基于 MF-R 和 AWS 密钥管理机制的物联网健康监测大数据分析系统 *

臧艳辉 [1]，赵雪章 [1]，席运江 [2]

(1. 佛山职业技术学院，广东 佛山 528137; 2. 华南理工大学，广州 510000)

摘　要：带有传感器的可穿戴式医疗设备不断生成大量数据，由于数据的复杂性，难以通过处理和分析大数据来找到有价值的决策信息。为了解决这个问题，提出了一种新的物联网体系结构，用于存储和处理医疗应用的可扩展传感器数据（大数据）。所提出的架构主要由两个子架构组成：Meta Fog 重定向（MF-R）架构和 AWS 密钥管理机制。MF-R 架构使用 Apache Pig 和 Apache HBase 等大数据技术来收集和存储不同传感器设备生成的传感器数据，并利用卡尔曼滤波消除噪声。AWS 密钥管理机制使用密钥管理方案，目的是保护云中的数据，防止未经授权的访问。当数据存储在云中时，所提出的系统能够使用随机梯度下降算法和逻辑回归来开发心脏病的预测模型。仿真实验表明，与其他几种算法相比，提出的算法具有更小的误差，且在吞吐量、准确度等方面具有一定的优越性。

关键词：无线传感器网络；物联网；大数据；卡尔曼滤波；云计算；AWS 密钥管理机制

中图分类号：TP391　　doi: 10.3969/j.issn.1001-3695.2018.05.0275

IoT health monitoring big data analysis system based on MF-R and AWS key management mechanism

Zang Yanhui[1], Zhao Xuezhang[1], Xi Yunjiang[2]

(1. *Foshan Polytechnic*, *Foshan Guangdong* 528137, *China*; 2. *South China University of Technology*, *Guangzhou* 510641, *China*)

Abstract: Wearable medical devices with sensor continuously generate enormous data, due to the complexity of the data, it is difficult to process and analyze the big data for finding valuable information that can be useful in decision-making. In order to overcome this issue, this paper proposed a new architecture for the implementation of IoT to store and process scalable sensor data (big data) for health care applications. The proposed architecture consisted of two main sub architectures, namely, Meta Fog-Redirection (MF-R) and AWS key management mechanism. MF-R architecture used big data technologies such as Apache Pig and Apache HBase for collection and storage of the sensor data generated from different sensor devices and it also used kalman filter for removal of noise. AWS key management mechanism used a key management scheme to protect data in the cloud and prevent unauthorized access. When data was stored in the cloud, the proposed system could use stochastic gradient descent algorithms and logistic regression to develop a predictive model of heart disease. Simulation experiments show that compared with other algorithms, the proposed algorithm has smaller error and it has certain advantages in terms of throughput and accuracy.

Key words: Wireless sensor network; Internet of Things; Big data; Kalman filter; Cloud computing; AWS key management mechanism

0　引言

物联网（Internet of things ，IOT）[1]是物理对象与用于收集和交换数据的网络连接而成的，物联网的发展使得传感器的应用越来越广泛[2]。例如，使用可穿戴式传感器装置，对患者进行周期性的连续生理监测，可以给出生理运动和饮食习惯建议。在此期间，可穿戴式传感器将持续获取患者的健康数据，并将其存储[3]，这有助于医生诊断患者的健康状况。在基于 IOT 的医疗健康应用中，“大数据”起着至关重要的作用[4]。现代医疗保健系统通常采用临床数据以增加在线可用的临床记录。此外，在 IOT 大数据分析中，通常借助于云计算[5]和雾计算[6]以提高效率，并减少需要从物理设备传输到云端的数据量[7,8]。科研人员针对 IOT 医疗系统做出了许多改进，Jee 和 Kim 描述了基于大数据分析的医疗系统改进方式，以改善医疗系统，选择适当

收稿日期：2018-05-04；**修回日期：**2018-06-29　**基金项目：**国家自然科学基金面上项目（71371077）；佛山市科技计划项目（2015AB004241）

作者简介：臧艳辉（1978-），女，湖北襄阳人，高级工程师，硕士，主要研究方向为大数据、图形图像处理等（zangshirley@126.com）；赵雪章（1972-），男，河南南阳人，副教授，硕士，主要研究方向为大数据、图像处理等；席运江（1973-），男，河南南阳人，副教授，博士，主要研究方向为大数据、知识网络等.

的治疗方式[9,10]。弗吉尼亚大学开发了 Alarm-Net 架构，用于监控患者健康状况[11]。该框架有三层架构，第一层使用 BP 传感器和 ECG（心电图）传感器监测患者的生理状况。第二层架构用于观察环境参数，如灰尘、热量、运动和光线等。第三层通过使用无线通信协议和网络架构将原始观测信号传输到目的地[12]。本文提出了一种新的物联网体系结构来存储和处理医疗应用的传感器数据。提出的架构包括 Meta Fog 重定向（MF-R）架构和 AWS 密钥管理机制两部分。其中，MF-R 架构主要用于收集和存储传感器数据，并利用卡尔曼滤波消除噪声。AWS 密钥管理机制使用密钥管理方案以保护云中的数据，防止未经授权的访问。

1 融合 MF-R 和 AWS 密钥管理机制的大数据分析系统

本文提出一种新的架构系统，用于存储和处理医疗保健应用中可扩展的传感器数据（大数据）。在图 1 中，描绘了各种设备物体之间的连接通信。

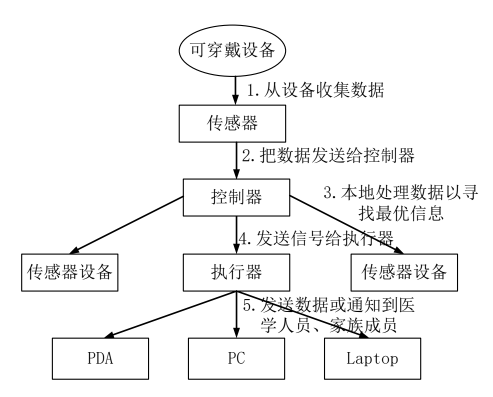


图 1 医疗物联网中物与物之间的通信

1.1 MF-R 架构

所提出的架构由三个不同阶段组成：数据收集阶段、数据处理传输阶段和大数据存储阶段。

a）数据收集阶段，用于持续监测患者的健康状况。因此，当呼吸频率、心率、血压、体温和血糖超过正常值时，设备通过雾计算使用无线网络向医生发送具有临床数据的警报消息。在数据收集阶段，通过雾计算收集的传感器数据与云连接，以便传输并存储到数据库中。

b）在数据处理传输阶段，所提出的架构使用“s3cmd utility”方法将临床数据移入 Amazon S3。医疗机构使用亚马逊云中的 EC2 来存储患者的传感器数据（生理数据）。但是，此传感器数据集在 Amazon EMR 中不可用，因为传感器数据仅存储在运行 EC2 实例的本地磁盘上。因此，必须把它转移到 EMR。Amazon EMR 在云中提供 Hadoop 服务来处理大数据[13]。

c）通常情况下，在 Amazon S3 中存储大量数据是不可能的[14]。为了解决这个问题，所提出的架构在数据存储阶段使用 Apache Pig 将大量的传感器数据（临床数据）从 Amazon S3 传输到 Apache HBase，如图 2 所示。

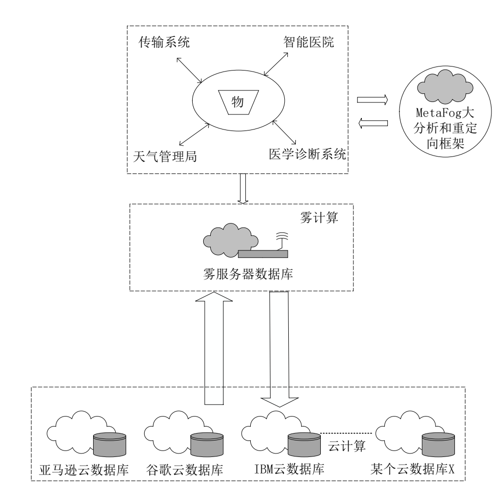


图 2 物联网大数据生态系统架构

1.1.1 数据收集阶段和处理传输阶段

将从雾计算中收集的传感器数据传输并存储到数据库中。所提出的架构将数据移入 Amazon S3，通过 Amazon EMR 访问。

本文通过应用数据过滤来提高数据处理效率，卡尔曼滤波器(Kalman filter，KF)用于在提出的框架中执行数据过滤[15]。KF 是一个最优估计器，它可以从感测数据中去除噪声。算法表示如下所示。它最初假定当前状态 f_k 是从先前状态 f_{k-1} 发展而来的。当前状态观察由 h_k 表示。$\hat{f}_{k|k-1}$ 表示时间 k 处的 f 的估计，而估计精度由 $G_{k|k-1}$ 表示。KF 有助于以最小的内存消耗进行处理。

KF 原理

1 初始化

T_k -状态转换模型(应用到上一状态 f_{k-1})

O_k -观察模型

Q_k -过程噪音的协方差

R_k -观察噪音的协方差

T_k -控制输入模型(应用到控制向量 v_k)

$w_k \sim (0, Q_k)$

2 使用上一状态 f_{k-1} 计算新的状态 f_k

$f_k = T_k f_{k-1} + C_k v_k + w_k$

$h_k = o_k f_k + u_k$

$u_k \sim (0, R_k)$

3 从上一状态得到的当前状态评估

评估的状态

$$\hat{f}_{k|k-1} = T_k \hat{f}_{k-1|k-1} + c_k v_k$$

预测的协方差

$$G_{k|k-1} = T_k G_{k-1|k-1} T_k^T + Q_k$$

```
联合当前预测和当前观测方程
    当前观测方程
```

$$\tilde{x}_k = h_k - O_k \hat{f}_{k|k-1}$$

```
    观测协方差方程
```

$$S_k = O_k G_{k|k-1} O_k^T + R_k$$

```
    最佳增益
```

$$K_k = G_{k|k-1} O_k^T S_k^{-1}$$

通常，Amazon EMR 处理存储在 Amazon S3 或 HDFS 存储中的数据。有必要为存储数据的 Amazon S3 提供全球唯一的名称。每个 Amazon S3 都有一个唯一的 URL 命名约束，来与其余 Amazon S3 交互。通常，有多种方法可将数据加载到 Amazon S3 中。本文所提出的架构使用“s3cmd utility”方法将数据移入 Amazon S3。S3cmd 负责上传、检索和管理云存储服务提供商的数据，这是一个适用于 S3 批处理脚本和自动备份的命令执行程序。

MetaFog 重定向架构

```
  处理审计日志文件，以检查登录用户的详细信息，比如他/她/ AI 机器的轨迹状态等。
选择患者编号 patient_id，密码 password，地理位置 geolocation，患
  者名字 patient_name，患者住址 patient_address，系统日期 sysdate;
如果有任何恶意用户试图登录应用程序
 执行
     插入更新威胁值 Threat_updated（patient_id，password，
       geolocation，patient_name，patient_address，sysdate）;
     返回-1;
否则
   列表 l=从 Meta 数据存储云中选择数据存储供应商名字、数据范
   围，其中登录应用程序的客户 Id =? ；
结束
如果数据存储供应商名字=“Fog_Sever_1”并且数据范围=“临界”）
 执行
     转到：Fog_Sever_DB_1
     从 Fog_Sever_DB_1 中选择*，其中登录应用程序的患者
       Id=?
 结束
如果数据存储供应商名字=“Fog_Sever_2”并且数据范围=“敏感”
  执行
     转到：Fog_Sever_DB_2
     从 Fog_Sever_DB_2 中选择*，其中：登录应用程序的患者
       Id=?
如果数据存储供应商名字=“Fog_Sever_3”并且数据范围=“正常”
  执行
     转到：Fog_Sever_DB_3
     从 Fog_Sever_DB_3 中选择*，其中：登录应用程序的患者
       Id =?
结束
```

接下来使用 mb 选项创建一个名为 clinical-emr 的新存储，使用“s3cmd put”将传感器数据移入 Amazon S3。

1.1.2 大数据存储阶段

Amazon EMR 提供了多种方法将大数据传送到集群中，可以将数据上传到 Amazon S3，然后使用 Amazon EMR 的内置功能将数据加载到 HBase 集群中。所提出的架构使用 Apache Pig 将传感器数据从 Amazon S3 传输到 HBase，HBase 是一个可扩展的分布式数据库，可以以分布式方式存储大量行和列，并可供集群中的所有节点使用，HBase 遵循列式数据库存储。Apache Pig 是分析大量数据的一个平台，将数据表示为数据流。Pig Latin 是一种脚本语言，用于 ETL（提取、转换和加载）过程以及对结构化、半结构化或非结构化数据的 ad hoc 数据分析。所提出的架构使用 Apache Pig 将新的增量临床数据与以前的数据结合到 HBase 中。

1.2　基于 AWS 密钥管理机制的安全服务

为了保证雾计算和云计算之间的安全服务，在本文提出的框架中使用了 AWS 密钥管理机制和 AWS Cloud Trail。AWS 密钥管理机制能够让用户安全、快捷的使用加密密钥。该机制在 Amazon S3、EBS 等 AWS 服务中已被广泛应用；此外，该机制还被应用在 AWS Cloud Trail 中，以方便用户更好的使用密钥管理服务，同时将用户日志记录下来并将其传输到 Amazon S3[16]。重定向体系结构如图 3 所示。所提出的安全框架，可防止在尝试登录到云应用程序时发生未经授权的访问。

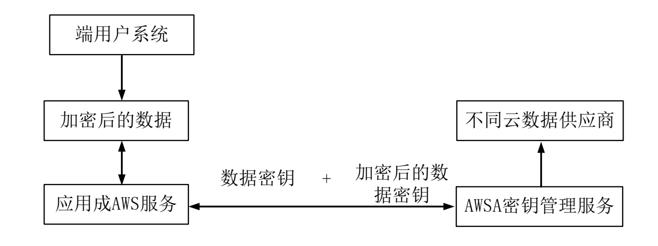


图 3　重定向体系安全框架

1.2.1 雾计算到云计算

科研人员把雾计算定义为云计算的延伸。使用雾计算改善的新兴网络范例，具有较少延迟和抖动。雾计算也被称为雾网络，用于分配云中计算资源和应用服务。雾计算的主要目标是提高可扩展性、效率并减少需要传输到云中进行数据分析、处理和存储的数据量。在雾计算环境中，数据处理总是在移动设备的数据中心或网关、智能路由器的网络边缘进行。

在本文体系结构中，可穿戴传感器设备将健康数据传输到雾服务器。雾服务器采用近边缘技术，将智能医疗应用中的健康监测设备连接起来。边缘计算在雾计算的安全性、保密性和系统可靠性方面发挥着至关重要的作用。近边缘技术的主要目标是减少传输每一比特通过云通道所需的带宽量。雾计算遵循网络边缘技术，因此从可穿戴医疗设备收集数据需要花费的时间更少了。在本文中，MetaFog 重定向结构根据数据分类（敏感、临界和正常）将健康数据重定向到不同的雾服务器。另外，该体系结构执行以下功能：安全性和日志文件处理，以防止来自恶意用户和未经授权用户的数据入侵。

1.2.2 把雾计算安全集成到云计算

云计算安全由策略和基于控制的技术组成，旨在保护与云计算相关的数据、应用程序、信息和基础架构。在本文所提出的架构中，当把雾集成到云环境时，安全性是主要问题。所提出的安全架构使用安全术语作为公钥和私钥、加密和解密，加密身份访问管理和 PKI 证书授权来保护云中的数据和应用程序。

1.2.3 逻辑回归随机梯度下降的 MapReduce 实现

所提出的物联网数据监测系统使用基于 MapReduce 的随机梯度下降与逻辑回归来开发预测模型。该预测模型使用 Apache HBase 中提供的数据。该逻辑回归使用现有数据来开发预测模型。使用 MapReduce 实现基于 SGD 的逻辑回归，如下所示：

输入：n 个临床记录集合 $V=\{(a_i\in R^d;b_i\in R)\}_{i=1}^n$；$\theta^t\in R^d$，为 t 时刻的权重；n 是学习速率。

输出：$\theta^{(t+1)}\in R^d$，为 $(t+1)$ 时刻的权重。

MAP 算法

class MAPPER
　method INITIALIZE
　double $\theta=1$;
　double n=0.1;
　#key k ;value v ;
　method MAP(string k ;double v);
　执行
　　$(a_i;b_i)=(k;v)$
　　r=rand() #number of reducer
　　EmithIntermediate $(r;(a_i;b_i))$;
　结束

Reduce 算法

class REDUCE
　#key k ;value v ;
　method REDUCE (string k ; double $v[1;2;\cdots;r]$)
　执行
　　$k^{\theta^{(t+1)}}=\theta^t$;
　　当 $(a_i;b_i)\in v[1;2;\cdots;r]$ 时，
　　　$k^{\theta^{(t+1)}}=k^{\theta^{(t+1)}}-n\nabla F\left(k^{\theta^{(t+1)}}\right)$;
　　结束
　返回(string k ; double $k^{\theta^{(t+1)}}$);
　结束
　n 个临床记录集合 $V=\{(a_i\in R^d;b_i\in R)\}_{i=1}^n$；
　每台机器接受 M 个存储块。

其中：$M=\frac{|V|}{m}$；在 MapReduce 框架中，M 由映射器自动分配[17]。最终权重 $\theta^{(t+1)}$ 和权重 $\theta^{(t+1)}$ 的平均值可以用 $\theta^{(t+1)}=\frac{1}{M}\sum_{k=1}^{M}k^{\theta^{(t+1)}}$ 来定义。

2　结果分析

把从各种传感器设备收集而来的数据分类为敏感数据、临界数据和正常数据，然后将分类结果存储在各种表格中（表 1~5）。假设患者的临床参数是临界数据，所有临界数据存储在表 1 中。表 2 描述了患者的个人数据，其中包括姓名、年龄、性别和地址；表 3 描绘了正常患者的数据，包括患者 ID、性别和年龄；表 4 描述了数据中心的详细信息，包括数据中心名称、位置及其类别（临界、敏感、正常）；表 5 描述了日志文件，并存储了与入侵者或未经授权访问有关的信息。

表 1　患者的临床参数

编号	患者临床参数	家/位置详细地址
C101	体温	波士顿
T202	心率	亚特兰大
S301	血糖	新德里

表 2　患者的个人资料

编号	名字	性别	年龄	位置
A231	Albert	男	47	那格浦尔
D342	David	女	58	班加罗尔
C352	Hendry	女	39	纽约

表 3　正常患者的数据

编号	性别	年龄	位置
A231	男	45	澳大利亚
D342	女	60	加拿大
C352	女	39	韩国

表 4　数据中心

数据存储供应商	数据中心位置	数据范围	患者编号
亚马逊	波士顿	临界	C101
谷歌	纽约	敏感	C352
微软	加拿大	正常	D342

表 5　有关入侵者的信息

编号	地理位置	IP 地址	日期和时间
C101	美国	10.10.142.107	12-1-2015: 14:23:03
T202	亚特兰大	12.11.112.108	14-1-2015: 14:23:03
D342	澳大利亚	11.10.122.182	12-1-2015: 14:43:08
X112	英国	22.11.111.107	22-1-2015: 14:23:05
N843	印度	12.10.146.162	14-1-2015: 15:23:07

根据不同临床参数值的 CPU 使用情况对所提出系统的性能进行分析。通过定量评估，获得从物联网传感器设备接收各种临床参数的时间间隔。图 4 表示从 IoT 传感器设备获得临床参数的不同数据集时，各种 CPU 的使用情况。图 5 表示使用 MapReduce 实现逻辑回归的性能评估。

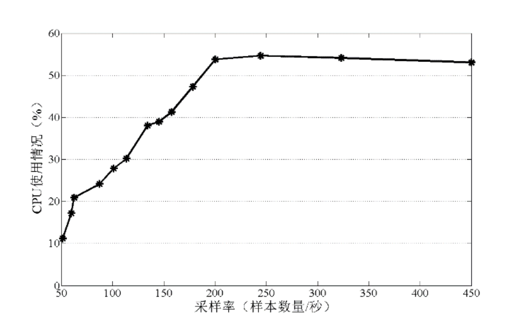


图 4　在不同临床参数采样率下的 CPU 使用情况

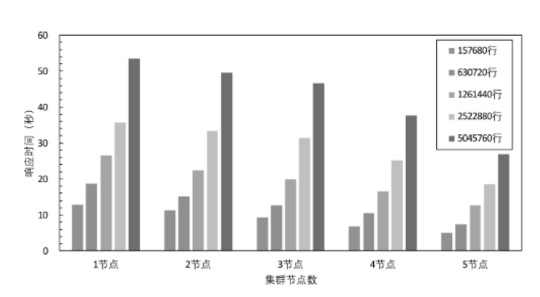


图 5　性能评估

3　性能评估

本文所提出的基于物联网的健康监测系统使用随机梯度下降算法和逻辑回归来开发心脏病的预测模型。从 CHDD 据库收集先前的临床记录和患者的传感器数据。在此过程中，健康监测系统通过云和大数据技术，使用由身体传感器设备获取的当前传感器数据。预测分类表如表 6 所示。所提出基于 MapReduce 的预测模型的性能分析如表 7 所示。所提出的心脏病预测模型能够对心脏病进行有效的分类，训练样本和验证样本的准确度分别为 74.39％和 73.92％。在灵敏度、特异度、精度、召回率和 F 均值这些指标的帮助下，评估预测模型的性能；这些指标的定义如下：

$$灵敏度=\frac{真正(TP)}{真正(TP)+假负(FN)} \tag{1}$$

$$特异度=\frac{真负(TN)}{假正(FP)+真负(TN)} \tag{2}$$

$$准确度=\frac{(TP)+(TN)}{(TP)+e(FN)+(FP)+(TN)} \tag{3}$$

$$精度=\frac{真正(TP)}{真正(TP)+假正(FP)} \tag{4}$$

$$召回率=\frac{真正(TP)}{真正(TP)+假负(FN)} \tag{5}$$

$$F均值=2*\frac{精度*灵敏度}{精度+灵敏度} \tag{6}$$

表 6　训练数据和测试数据

训练样本				验证样本			
	Yes（1）	No（0）	总计		Yes（1）	No（0）	总计
Yes（1）	89	33	122	Yes（1）	36	11	47
No（0）	25	64	89	No（0）	14	31	45

表 7　性能评估

训练样本					
准确度	灵敏度	特异度	精度	召回率	F 均值
74.39%	78.07%	66.13%	72.95%	79.22%	75.42%
验证样本					
准确度	灵敏度	特异度	精度	召回率	F 均值
73.92%	71.82%	74.2%	76.59%	72.74%	72.92%

把本文算法的结果与文献[18~20]的算法的结果进行比较分析，如表 8 所示。计算每个机器学习算法的均方误差(MSE)、均方根误差（RMSE）、R 平方、Q 平方、误差平方和（SSE）、误差平方和总和（TSSE）、平均绝对误差（MAE）、平均绝对误差百分比（MAPE）和均方差（MSD），部分结果如图 6 所示。可以看出，与其他算法相比，本文算法的性能更好。

MSE 是估计值和测量值之间的差，其计算公式如下：

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2 \tag{7}$$

其中：n 是观测次数，y_i 是观测值，$\hat{y}_i$ 是预测值。

RMSE 表示测量模型或估计器预测值与观测值之间的差，均方根误差（RMSE）是 MSE 的平方根：

$$RMSE=\sqrt{MSE} \tag{8}$$

R^2 是数据与拟合回归线接近度的统计测量，R^2 的定义如下：

$$R^2=1-\frac{SSE}{TSS} \tag{9}$$

其中，SSE 是每个观测值与其组平均值的误差平方之和：

$$SSE=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2 \tag{10}$$

SSE 同样可以表示为

$$SSE=n\times MSE \tag{11}$$

TSS 为所有观测数据中，每个观测值与总平均值的误差平方之总和。TSS 公式如下：

$$TSS=\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2 \tag{12}$$

其中：$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$

因此，R^2 也被定义为

$$R^2 = 1 - \frac{n \times MSE}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{13}$$

Q^2 是 MSE 与响应变量 Y 的方差之比。Q^2 公式如下：

$$Q^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{14}$$

R^2 由下式定义：

$$R^2 \sim 1 - Q^2 \tag{15}$$

平均绝对误差（MAE）衡量测试样本中，预测与实际观测值间绝对差的平均值，其中所有个体差异具有相同的权重。MAE 公式如下：

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{16}$$

平均绝对误差百分比（MAPE）以百分比形式衡量误差的大小。MAPE 被定义为

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{17}$$

使用均方差（MSD）测量拟合时间序列值的精度。MSD 的计算方法为

$$MSD = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|^2 \tag{18}$$

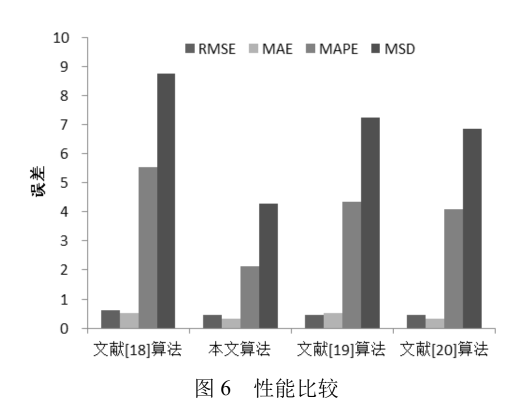


图 6　性能比较

吞吐量是根据在给定时间段内，从物联网传感器设备发送到传感器服务器的数据量来计算的。例如，如果在 10 s 内的数据传输量为 16 个字节，则物联网传感器设备的吞吐量为每秒 1.6 个字节。图 7 显示了使用四种算法测量得到的各种物联网健康监测设备的吞吐量，由图可知，本文算法的吞吐量相对于其余算法总是最大，其性能较好。

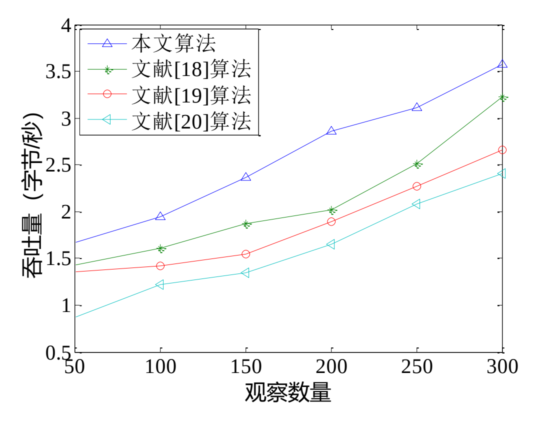


图 7　吞吐量比较

表 8　性能分析

不同算法	性能指标								
	RMSE	MSE	SSE	TSS	R^2	Q^2	MAE	MAPE	MSD
文献[18]算法	0.629	0.175	10.764	13.864	0.307	0.892	0.503	5.395	8.742
本文算法	0.441	0.078	8.424	12.024	0.299	0.705	0.315	2.143	4.257
文献[19]算法	0.448	0.141	12.984	16.884	0.223	0.786	0.511	4.345	7.241
文献[20]算法	0.445	0.104	11.267	14.867	0.242	0.757	0.326	4.092	6.844

4　结束语

本文提出了一种新的物联网体系结构以处理医疗应用中的可扩展传感器数据。所提出的架构主要由两个子架构组成：Meta Fog 重定向（MF-R）架构和分组 AWS 密钥管理机制。MF-R 架构使用 Apache Pig 和 Apache HBase 等大数据技术来收集和存储不同传感器设备生成的传感器数据，并利用卡尔曼滤波消除噪声。AWS 密钥管理机制用于确保雾计算与云计算的集成。该体系结构还使用密钥管理服务和数据分类函数来提供安全服务，并使用基于 MapReduce 的预测模型来预测心脏病。本文通过计算吞吐量、灵敏度、准确度和 F-measure 等性能评估参数证明了所提出的架构以及预测模型的有效性。

参考文献：

[1] Xia Feng, Yang Tianruo L, Wang Lizhe, *et al.* Internet of Things [J]. International Journal of Communication Systems, 2012, 25 (9): 1101-1102.

[2] 丁治明，高需．面向物联网海量传感器采样数据管理的数据库集群系统框架 [J]．计算机学报, 2012, 35 (6): 1175-1191. (Ding Zhiming, Gao Xu. A database cluster system framework for managing massive sensor sampling data in the Internet of things [J]. Chinese Journal of Computers, 2012, 35 (6): 1175-1191.)

[3] Lorincz K, Malan D, Fulford J T, *et al.* Sensor networks for emergency response: challenges and opportunities [J]. IEEE Pervasive Computing, 2005, 3 (4): 16-23.

[4] 张引，陈敏，廖小飞．大数据应用的现状与展望 [J]．计算机研究与发展, 2013, 50 (s2): 216-233. (Zhang Yin, Chen Min, Liao Xiaofei. Big data applications: a survey [J]. Journal of Computer Research and Development, 2013, 50 (s2): 216-233.)

[5] 李慧芳，刘秀平．凸价格函数下基于堆栈执行的云计算资源调度方案 [J]．计算机应用研究, 2017, 34 (10): 3129-3132. (Li Huifang, Liu Xiuping. Resource scheduling scheme based on stack process under convex price function in cloud computing [J]. Application Research of Computers, 2017, 34 (10): 3129-3132.)

[6] Bonomi F, Milito R, Zhu Jiang, *et al.* Fog computing and its role in the Internet of things [C]// Proc of the 1st Edition of the MCC Workshop on Mobile Cloud Computing. New York: ACM Press, 2012: 13-16.

[7] Li Shancang, Xu Lida, Wang Xinheng. Compressed sensing signal and data acquisition in wireless sensor networks and Internet of things [J]. IEEE Trans on Industrial Informatics, 2013, 9 (4): 2177-2186.

[8] 何秀丽，任智源，史晨华，等．面向医疗大数据的云雾网络及其分布式计算方案 [J]．西安交通大学学报, 2016, 50 (10): 71-77. (He Xiuli, Ren Zhiyuan, Shi Chenhua, *et al.* A cloud and fog network architecture for medical big data and its distributed computing scheme [J]. Journal of Xi'an Jiaotong University, 2016, 50 (10): 71-77.)

[9] Jee K, Kim G H. Potentiality of big data in the medical sector: focus on how to reshape the healthcare system [J]. Healthcare Informatics Research, 2013, 19 (2): 79–85.

[10] Chawla N V, Davis D A. Bringing big data to personalized healthcare: a patient-centered framework [J]. Journal of General Internal Medicine, 2013, 28 (3): 660-665.

[11] Masdari M, Ahmadzadeh S. Comprehensive analysis of the authentication methods in wireless body area networks [J]. Security & Communication Networks, 2016, 9 (17): 4777-4803.

[12] 龙昭华，龚俊，王波，等．无线传感器网络中分簇安全路由协议保密通信方法的能效研究 [J]．电子与信息学报, 2015, 37 (8): 2000-2006. (Long Zhaohua, Gong Jun, Wang Bo, *et al.* Energy efficiency study of secret communication method on clustering secure routing in WSN [J]. Journal of Electronics & Information Technology, 2015, 37 (8): 2000-2006.)

[13] He Hui, Du Zhonghui, Zhang Weizhe, *et al.* Optimization strategy of Hadoop small file storage for big data in healthcare [J]. Journal of Supercomputing, 2016, 72 (10): 3696-3707.

[14] Santos M Y, Martinho B, Costa C. Modelling and implementing big data warehouses for decision support [J]. Journal of Management Analytics, 2017, 4 (2): 111-129.

[15] 胡振涛，胡玉梅，刘先省．量测提升卡尔曼滤波 [J]．电子学报, 2016, 44 (5): 1149-1155. (Hu Zhentao, Hu Yumei, Liu Xianxing. Kalman filter based on measurement lifting strategy [J]. Acta Electronica Sinica, 2016, 44 (5): 1149-1155.)

[16] Manogaran G, Thota C, Kumar V. Meta cloud data storage architecture for big data security in cloud computing [J]. Procedia Computer Science, 2016, 87 (3): 128-133.

[17] Kang D, Lim W S, Shin K, *et al.* Data//feature distributed stochastic coordinate descent for logistic regression [C]// Proc of ACM International Conference on Conference on Information and Knowledge Management. 2014: 1269-1278.

[18] Rehab M A, Boufarès F. Scalable massively parallel learning of multiple linear regression algorithm with MapReduce [C]// Proc of IEEE Trustcom//BigDataSE//ISPA. Washington DC: IEEE Computer Society, 2015: 41-47.

[19] Rejab F B, Nouira K, Trabelsi A. Real time SVM for health monitoring system [J]. Brain Informatics and Health, 2014, 2014 (8609): 301-312.

[20] Wang Limin, Chen Jiaxu, Fan Min, *et al.* Application of random forest data mining method to the feature selection for female sub-health state [C]// Proc of IEEE International Conference on Bioinformatics and Biomedicine Workshops. 2011: 651-654.